

## 분류 개념을 포함한 관속식물 목-과 부호 제안

송준호 선임연구원, 양선규 기술연구원, 최고야 선임연구원\*

한국한의학연구원 한약자원연구센터

### Suggestions for a code of orders–families in vascular plants including taxonomic concept

Jun-Ho Song, Sungyu Yang, Goya Choi\*

Herbal Medicine Resources Research Center, Korean Institute of Oriental Medicine

#### Abstract

The various database management system has been developed for storing, processing, and retrieving taxonomic data. In this paper, we present an order-family code, which is decimals, an 11-digit number. The order-family code contains phylogenetic relationships between each order and family and to help their database establishing effectively. To present the family name with code, we can intuitively know the systematic relationship and relative distance of a particular group. Moreover, this code could be applied in various ways for databases such as add a family code field, visualization of cladograms, etc. Thus, this suggested code presents new perspectives for future work on the effective establishing and managing of the biological database.

**Keywords:** Taxonomic database, Order, Family, Hierarchical level, APG system, Bioinformatics

#### 서론

생물정보학(bioinformatics)은 컴퓨터공학, 통계학 및 공학적 개념과 방법을 이용하여 생물정보를 연구하는 학문으로 크게 생물정보관리(biological information management)와 컴퓨터 생명공학(computational biology) 두 종류로 구분된다<sup>1,2)</sup>.

**Correspondence:** 최고야(Goya Choi)

111 Geonjae-ro, Naju-si, Jeollanam-do, Rep. of Korea

Tel: +82-61-338-7118, E-mail: serparas@kiom.re.kr

Received 2020-07-06, revised 2020-07-09, accepted 2020-07-14, available online 2020-07-16

doi:10.22674/KHMI-8-2-1



최근에는 생물정보학 내 생물정보관리의 관점에서 데이터의 수집(data capture), 보관(storage), 검색(retrieving)을 포함한 다양한 데이터 분류정보시스템과 데이터 처리(data processing) 기법이 개발되고 있다. 이러한 시스템은 생물지리학, 서지학 및 문헌적 데이터뿐만 아니라 계층적 생물학 계급에 해당되는 분류군의 명명법적, 분류학적 기재를 포함하는 다양한 생물 데이터베이스에 적용되고 있다<sup>3)</sup>. 특히, 분류 검색엔진(taxonomic search engine)을 기반으로 한 다음의 데이터베이스는 주로 트리 구조의 계층을 이루는 분류방식을 채택하고 있다: The Integrated Taxonomic Information System (ITIS)<sup>4)</sup>, The International Plant Names Index (IPNI)<sup>5)</sup>, The Plant List<sup>6)</sup>, Index Fungorum<sup>7)</sup>, The Universal Biological Indexer and Organizer (uBio)<sup>8)</sup>, The NCBI Taxonomy database<sup>9)</sup>, The Barcode of Life Data System (BOLD or BOLDSystems)<sup>10)</sup>.

이 방식은 생물분류체계를 일목요연하게 알 수 있다는 장점이 있으나, 전문 생물학 데이터베이스가 아닌 과(family) 또는 목(order) 수준의 정보만을 필요로 하는 경우 필요한 정보량에 비해 구현해야 할 알고리즘이 지나치게 복잡해진다는 단점이 있다. 예를 들어, 500 여 종의 약용식물 정보 제공을 위한 웹서비스를 만드는 경우, 계층적 계급인 역(domain) - 계(kingdom) - 문(phylum/division) - 강(class) - 목(order) - 과(family) - 속(genus) - 종(species)에 이르는 모든 트리 구조를 보여주는 것은 일반적으로 불필요하다. 약용(한약·생약) 자원의 기원종을 수재하고 있는 전 세계 모든 약전에서도 기원종의 분류체계는 '과' 수준에서만 제시하는 것으로 알려져 있다. 이렇게 단순히 일정 계급 수준에서만 정보를 제시하는 경우에는 데이터베이스 자원 소비를 최소화한다는 장점이 있으나, 각 과 사이의 유연관계에 대한 정보는 전혀 알 수 없다는 단점이 있다.

이에 약전 수재 기원종을 비롯한 한약·생약자원의 데이터베이스 구축을 위해 상기의 두 가지 방식을 절충하여, 각 과 사이의 유연관계를 포함하면서 데이터베이스 자원도 적게 소비하는 '목-과 부호'(order-family code)를 제안하고자 한다.

## 본론

### 1. 관속식물 목-과 부호

본 연구는 관속식물 분류군(Vascular plants; Tracheophytes)을 대상으로 수행하였으며, Angiosperm Phylogeny Group IV (APG IV)<sup>11)</sup>를 기반으로 한 Angiosperm Phylogeny Website, Version 14 (이하 AP Web)<sup>12)</sup> (2020. 07. 기준)에 제시된 목과 과를 목록화하여 부호를 부여하였다. 한편, 석송류(lycophyte) 및 양치류(ferns)의 경우, AP Web 을 기반으로 Pteridophyte Phylogeny Group (PPG)<sup>13)</sup>에 제시된 최신 분류체계를 반영하여 정리하였다. AP Web 에 제시된 분계도를 기반으로 '관속식물'에서 출발하여, 한 번 차상분지를 할 때마다 각 분지에 0 과 1 을 부여하고, 매 분지마다 자릿수를 추가하여 2 진수로 숫자를 부여하였다. 이를 목 단위가 나올 때까지 반복하며, 먼저 분기가 완료된 목은 마지막 목과 동일한 자릿수가 되도록 0 을 채워 정리하였다. 따라서 목 계급의 관속식물 분계도 내 가장 기저에 위치하며 첫 번째로 분기되는 석송목(Lycopodiales)의 경우 '000000000000000000000000'이 되고, 마지막으로 분기되는 산토끼꽃목(Dipsacales)의 경우 '1111111111111111111111111111'이 된다. 하지만 이는 27 자릿수로 너무 긴 부호이므로 10 진수로 변환하여 각각 '00000000'과 '134217727'로 표기하였다. 자릿수를 더 줄이기 위해서는 16 진수로 각각 '000000', '7fffff'로 표기할 수 있으나 7 자리 16 진수보다는 9 자리 10 진수가 데이터의 관리에 효율적이고 편리하므로 10 진수 9 자리 부호를 관속식물 목의 고유부호로 부여하였다.



과의 부호 역시 각 목에서 동일한 방식을 취하고자 하였으나, 아직 과 수준에서 정확히 차상으로 분지하는 계통학적 연구 결과가 이루어지지 않아<sup>11)</sup> 2 진수의 부호를 부여할 수 없었다. 따라서 하나의 목에 속해 있는 각 과는 00 부터 시작하는 두 자리 10 진수의 일련번호를 부여하되, 분계도에서 먼저 분지되는 과보다 나중에 분지되는 과에 더 큰 숫자를 부여하였으며, 다중 분기(polytomy)의 경우 AP Web 에 제시된 순으로 숫자를 부여하여 표기하였다. 예컨대 석죽목(Caryophyllales, 134216704) 안에서 끈끈이귀개과(Droseraceae)는 '00', 선인장과(Cactaceae)는 '37'이 된다. 한 목 아래에 가장 많은 과가 있는 경우는 석죽목(37 개 과)이므로, 각 과의 일련번호는 두 자리로 설정하였다. 따라서 최종 과 부호는 목 부호와 과 일련번호를 하이픈으로 이어서 총 12 자리의 부호로 제시하였다. (표 1)

표 1. 관속식물 목-과 부호

목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
Lycopodiales	00000000000000000000000000000000	000000000	Lycopodiaceae	00	000000000-00
Isoëtales	01000000000000000000000000000000	033554432	Isoëtaceae	00	033554432-00
Selaginellales	01100000000000000000000000000000	050331648	Selaginellaceae	00	050331648-00
Equisetales	10000000000000000000000000000000	067108864	Equisetaceae	00	067108864-00
Psilotales	10100000000000000000000000000000	083886080	Psilotaceae	00	083886080-00
Ophioglossales	10101000000000000000000000000000	088080384	Ophioglossaceae	00	088080384-00
Marattiales	10110000000000000000000000000000	092274688	Marattiaceae	00	092274688-00
Osmundales	10111000000000000000000000000000	096468992	Osmundaceae	00	096468992-00
Hymenophyllales	10111100000000000000000000000000	098566144	Hymenophyllaceae	00	098566144-00
Gleicheniales	10111110000000000000000000000000	099614720	Matoniaceae	00	099614720-00
			Dipteridaceae	01	099614720-01
			Gleicheniaceae	02	099614720-02
Schizaeales	10111111000000000000000000000000	100139008	Lygodiaceae	00	100139008-00
			Schizaeaceae	01	100139008-01
			Anemiaceae	02	100139008-02
Salviniales	10111111100000000000000000000000	100401152	Salviniaceae	00	100401152-00
			Marsileaceae	01	100401152-01
Cyatheaales	10111111110000000000000000000000	100532224	Thyrsopteridaceae	00	100532224-00
			Loxsomataceae	01	100532224-01
			Culcitaceae	02	100532224-02
			Plagiogyriaceae	03	100532224-03
			Cibotiaceae	04	100532224-04
			Metaxyaceae	05	100532224-05
			Dicksoniaceae	06	100532224-06
Cyatheaceae	07	100532224-07			
Polypodiales	10111111111000000000000000000000	100597760	Saccolomataceae	00	100597760-00
			Cystodiaceae	01	100597760-01
			Lonchitidaceae	02	100597760-02
			Lindsaeaceae	03	100597760-03
			Pteridaceae	04	100597760-04



목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
			Dennstaedtiaceae	05	100597760-05
			Cystopteridaceae	06	100597760-06
			Rhachidosoraceae	07	100597760-07
			Diplaziopsidaceae	08	100597760-08
			Desmophlebiaceae	09	100597760-09
			Hemidictyaceae	10	100597760-10
			Aspleniaceae	11	100597760-11
			Woodsiaceae	12	100597760-12
			Onocleaceae	13	100597760-13
			Blechnaceae	14	100597760-14
			Athyriaceae	15	100597760-15
			Thelypteridaceae	16	100597760-16
			Didymochlaenaceae	17	100597760-17
			Hypodematiaceae	18	100597760-18
			Dryopteridaceae	19	100597760-19
			Nephrolepidaceae	20	100597760-20
			Lomariopsidaceae	21	100597760-21
			Tectariaceae	22	100597760-22
			Oleandraceae	23	100597760-23
			Davalliaceae	24	100597760-24
			Polypodiaceae	25	100597760-25
Cycadales	11000000000000000000000000000000	100663296	Cycadaceae	00	100663296-00
			Zamiaceae	01	100663296-01
Ginkgoales	11001000000000000000000000000000	104857600	Ginkgoaceae	00	104857600-00
Cupressales	11010000000000000000000000000000	109051904	Araucariaceae	00	109051904-00
			Podocarpaceae	01	109051904-01
			Sciadopityaceae	02	109051904-02
			Taxaceae	03	109051904-03
			Cupressaceae	04	109051904-04
Pinales	11011000000000000000000000000000	113246208	Pinaceae	00	113246208-00
Gnetales	11011100000000000000000000000000	115343360	Ephedraceae	00	115343360-00
			Gnetaceae	01	115343360-01
			Welwitschiaceae	02	115343360-02
Amborellales	11100000000000000000000000000000	117440512	Amborellaceae	00	117440512-00
Nymphaeales	11110000000000000000000000000000	125829120	Hydatellaceae	00	125829120-00
			Cabombaceae	01	125829120-01
			Nymphaeaceae	02	125829120-02
Austobaileyales	11111000000000000000000000000000	130023424	Austrobaileyaceae	00	130023424-00
			Schisandraceae	01	130023424-01
			Trimeniaceae	02	130023424-02



목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
Chloranthales	11111100000000000000000000	132120576	Chloranthaceae	00	132120576-00
Magnoliales	11111101000000000000000000	132644864	Myristicaceae	00	132644864-00
			Magnoliaceae	01	132644864-01
			Degeneriaceae	02	132644864-02
			Himantandraceae	03	132644864-03
			Eupomatiaceae	04	132644864-04
			Annonaceae	05	132644864-05
Laurales	11111101010000000000000000	132775936	Calycanthaceae	00	132775936-00
			Siparunaceae	01	132775936-01
			Gomortegaceae	02	132775936-02
			Atherospermataceae	03	132775936-03
			Monimiaceae	04	132775936-04
			Hernandiaceae	05	132775936-05
			Lauraceae	06	132775936-06
Canellales	11111101100000000000000000	132907008	Canellaceae	00	132907008-00
			Winteraceae	01	132907008-01
Piperales	11111101110000000000000000	133038080	Aristolochiaceae	00	133038080-00
			Piperaceae	01	133038080-01
			Saururaceae	02	133038080-02
Acorales	11111110000000000000000000	133169152	Acoraceae	00	133169152-00
Alismatales	11111101000000000000000000	133431296	Araceae	00	133431296-00
			Tofieldiaceae	01	133431296-01
			Hydrocharitaceae	02	133431296-02
			Butomaceae	03	133431296-03
			Alismataceae	04	133431296-04
			Scheuchzeriaceae	05	133431296-05
			Aponogetonaceae	06	133431296-06
			Juncaginaceae	07	133431296-07
			Maundiaceae	08	133431296-08
			Posidoniaceae	09	133431296-09
			Ruppiaceae	10	133431296-10
			Cymodoceaceae	11	133431296-11
			Zosteraceae	12	133431296-12
Potamogetonaceae	13	133431296-13			
Petrosaviales	11111101100000000000000000	133562368	Petrosaviaceae	00	133562368-00
Dioscoreales	11111101110000000000000000	133627904	Nartheciaceae	00	133627904-00
			Taccaceae	01	133627904-01
			Thismiaceae	02	133627904-02
			Burmanniaceae	03	133627904-03
			Dioscoreaceae	04	133627904-04



목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
Pandanales	11111101110100000000000000	133644288	Velloziaceae	00	133644288-00
			Triuridaceae	01	133644288-01
			Stemonaceae	02	133644288-02
			Pandanaceae	03	133644288-03
			Cyclanthaceae	04	133644288-04
Liliales	11111101111000000000000000	133660672	Corsiaceae	00	133660672-00
			Campynemataceae	01	133660672-01
			Melanthiaceae	02	133660672-02
			Petermanniaceae	03	133660672-03
			Colchicaceae	04	133660672-04
			Alstroemeriaceae	05	133660672-05
			Ripogonaceae	06	133660672-06
			Philesiaceae	07	133660672-07
			Smilacaceae	08	133660672-08
Liliaceae	09	133660672-09			
Asparagales	11111101111100000000000000	133677056	Orchidaceae	00	133677056-00
			Boryaceae	01	133677056-01
			Blandfordiaceae	02	133677056-02
			Lanariaceae	03	133677056-03
			Asteliaceae	04	133677056-04
			Hypoxidaceae	05	133677056-05
			Ixioliriaceae	06	133677056-06
			Tecophilaeaceae	07	133677056-07
			Doryanthaceae	08	133677056-08
			Iridaceae	09	133677056-09
			Xeronemataceae	10	133677056-10
			Asphodelaceae	11	133677056-11
			Amaryllidaceae	12	133677056-12
Asparagaceae	13	133677056-13			
Arecales	11111101111110000000000000	133685248	Dasyopogonaceae	00	133685248-00
			Arecaceae	01	133685248-01
Poales	11111101111110000000000000	133689344	Typhaceae	00	133689344-00
			Bromeliaceae	01	133689344-01
			Rapateaceae	02	133689344-02
			Xyridaceae	03	133689344-03
			Eriocaulaceae	04	133689344-04
			Mayacaceae	05	133689344-05
			Thurniaceae	06	133689344-06
			Juncaceae	07	133689344-07
Cyperaceae	08	133689344-08			



목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
			Restionaceae	09	133689344-09
			Flagellariaceae	10	133689344-10
			Joinvilleaceae	11	133689344-11
			Ecdeiocoleaceae	12	133689344-12
			Poaceae	13	133689344-13
Commelinales	11111101111111100000000000	133691392	Commelinaceae	00	133691392-00
			Hanguanaceae	01	133691392-01
			Philydraceae	02	133691392-02
			Haemodoraceae	03	133691392-03
			Pontederiaceae	04	133691392-04
Zingiberales	11111101111111110000000000	133692416	Musaceae	00	133692416-00
			Heliconiaceae	01	133692416-01
			Strelitziaceae	02	133692416-02
			Lowiaceae	03	133692416-03
			Cannaceae	04	133692416-04
			Marantaceae	05	133692416-05
			Zingiberaceae	06	133692416-06
			Costaceae	07	133692416-07
Ceratophyllales	11111110000000000000000000	133693440	Ceratophyllaceae	00	133693440-00
Ranunculales	11111111000000000000000000	133955584	Eupteleaceae	00	133955584-00
			Lardizabalaceae	01	133955584-01
			Circaeasteraceae	02	133955584-02
			Menispermaceae	03	133955584-03
			Berberidaceae	04	133955584-04
			Ranunculaceae	05	133955584-05
			Papaveraceae	06	133955584-06
Proteales	111111110000000000000000	134086656	Sabiaceae	00	134086656-00
			Nelumbonaceae	01	134086656-01
			Platanaceae	02	134086656-02
			Proteaceae	03	134086656-03
Trochodendrales	111111111000000000000000	134152192	Trochodendraceae	00	134152192-00
Buxales	111111111100000000000000	134184960	Buxaceae	00	134184960-00
Gunnerales	111111111110000000000000	134201344	Gunneraceae	00	134201344-00
			Myrothamnaceae	01	134201344-01
Dilleniales	111111111111000000000000	134209536	Dilleniaceae	00	134209536-00
Saxifragales	111111111111010000000000	134211584	Peridiscaceae	00	134211584-00
			Paeoniaceae	01	134211584-01
			Altingiaceae	02	134211584-02
			Hamamelidaceae	03	134211584-03
			Cercidiphyllaceae	04	134211584-04



목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
			Daphniphyllaceae	05	134211584-05
			Crassulaceae	06	134211584-06
			Aphanopetalaceae	07	134211584-07
			Tetracarpaeaceae	08	134211584-08
			Penthoraceae	09	134211584-09
			Haloragaceae	10	134211584-10
			Iteaceae	11	134211584-11
			Grossulariaceae	12	134211584-12
			Saxifragaceae	13	134211584-13
			Cynomoriaceae	14	134211584-14
Vitales	11111111111110110000000000	134212608	Vitaceae	00	134212608-00
Zygophyllales	11111111111110110000000000	134213120	Krameriaceae	00	134213120-00
			Zygophyllaceae	01	134213120-01
Celastrales	11111111111110110100000000	134213248	Lepidobotryaceae	00	134213248-00
			Celastraceae	01	134213248-01
Oxalidales	11111111111110110101000000	134213280	Huaceae	00	134213280-00
			Connaraceae	01	134213280-01
			Oxalidaceae	02	134213280-02
			Cunoniaceae	03	134213280-03
			Brunelliaceae	04	134213280-04
			Cephalotaceae	05	134213280-05
			Elaeocarpaceae	06	134213280-06
Malpighiales	111111111111101101010000	134213296	Ctenolophonaceae	00	134213296-00
			Erythroxylaceae	01	134213296-01
			Rhizophoraceae	02	134213296-02
			Irvingiaceae	03	134213296-03
			Pandaceae	04	134213296-04
			Ochnaceae	05	134213296-05
			Bonnetiaceae	06	134213296-06
			Clusiaceae	07	134213296-07
			Calophyllaceae	08	134213296-08
			Hypericaceae	09	134213296-09
			Podostemaceae	10	134213296-10
			Lophopyxidaceae	11	134213296-11
			Putranjivaceae	12	134213296-12
			Caryocaraceae	13	134213296-13
			Centroplacaceae	14	134213296-14
			Elatinaceae	15	134213296-15
			Malpighiaceae	16	134213296-16
			Balanopaceae	17	134213296-17





목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
			Trigoniaceae	18	134213296-18
			Dichapetalaceae	19	134213296-19
			Euphroniaceae	20	134213296-20
			Chrysobalanaceae	21	134213296-21
			Humiriaceae	22	134213296-22
			Achariaceae	23	134213296-23
			Goupiaceae	24	134213296-24
			Violaceae	25	134213296-25
			Passifloraceae	26	134213296-26
			Lacistemataceae	27	134213296-27
			Salicaceae	28	134213296-28
			Peraceae	29	134213296-29
			Rafflesiaceae	30	134213296-30
			Euphorbiaceae	31	134213296-31
			Phyllanthaceae	32	134213296-32
			Picrodendraceae	33	134213296-33
Linaceae	34	134213296-34			
Ixonanthaceae	35	134213296-35			
Fabales	111111111111101101100000	134213312	Quillajaceae	00	134213312-00
			Fabaceae	01	134213312-01
			Surianaceae	02	134213312-02
			Polygalaceae	03	134213312-03
Rosales	111111111111101101100000	134213344	Rosaceae	00	134213344-00
			Barbeyaceae	01	134213344-01
			Dirachmaceae	02	134213344-02
			Rhamnaceae	03	134213344-03
			Elaeagnaceae	04	134213344-04
			Ulmaceae	05	134213344-05
			Cannabaceae	06	134213344-06
			Moraceae	07	134213344-07
Urticaceae	08	134213344-08			
Cucurbitales	111111111111101101110000	134213360	Anisophylleaceae	00	134213360-00
			Corynocarpaceae	01	134213360-01
			Coriariaceae	02	134213360-02
			Cucurbitaceae	03	134213360-03
			Tetramelaceae	04	134213360-04
			Daticaceae	05	134213360-05
			Begoniaceae	06	134213360-06
Apodanthaceae	07	134213360-07			
Fagales	111111111111101101111000	134213368	Nothofagaceae	00	134213368-00



목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
			Fagaceae	01	134213368-01
			Myricaceae	02	134213368-02
			Juglandaceae	03	134213368-03
			Ticodendraceae	04	134213368-04
			Betulaceae	05	134213368-05
			Casuarinaceae	06	134213368-06
Geraniales	1111111111111011110000000	134213376	Geraniaceae	00	134213376-00
			Francoaceae	01	134213376-01
Myrtales	1111111111111011110100000	134213440	Combretaceae	00	134213440-00
			Onagraceae	01	134213440-01
			Lythraceae	02	134213440-02
			Myrtaceae	03	134213440-03
			Vochysiaceae	04	134213440-04
			Melastomataceae	05	134213440-05
			Crypteroniaceae	06	134213440-06
			Alzateaceae	07	134213440-07
Penaecaceae	08	134213440-08			
Crossosomatales	1111111111111011110000000	134213504	Staphyleaceae	00	134213504-00
			Guamatelaceae	01	134213504-01
			Stachyuraceae	02	134213504-02
			Crossosomataceae	03	134213504-03
			Aphloiaceae	04	134213504-04
			Geissolomataceae	05	134213504-05
Strasburgeriaceae	06	134213504-06			
Picramniales	1111111111111011111000000	134213568	Picramniaceae	00	134213568-00
Sapindales	1111111111111011111100000	134213600	Biebersteiniaceae	00	134213600-00
			Nitrariaceae	01	134213600-01
			Kirkiaceae	02	134213600-02
			Burseraceae	03	134213600-03
			Anacardiaceae	04	134213600-04
			Sapindaceae	05	134213600-05
			Simaroubaceae	06	134213600-06
			Meliaceae	07	134213600-07
Rutaceae	08	134213600-08			
Huerteales	1111111111111011111110000	134213616	Gerrardinaceae	00	134213616-00
			Petenaecaceae	01	134213616-01
			Tapisciaceae	02	134213616-02
			Dipentodontaceae	03	134213616-03
Malvales	1111111111111011111111000	134213624	Neuradaceae	00	134213624-00
			Thymelaeaceae	01	134213624-01



목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
			Sphaerosepalaceae	02	134213624-02
			Bixaceae	03	134213624-03
			Cistaceae	04	134213624-04
			Sarcolaenaceae	05	134213624-05
			Dipterocarpaceae	06	134213624-06
			Cytinaceae	07	134213624-07
			Muntingiaceae	08	134213624-08
			Malvaceae	09	134213624-09
			Brassicales	1111111111111011111111100	134213628
Tropaeolaceae	01	134213628-01			
Moringaceae	02	134213628-02			
Caricaceae	03	134213628-03			
Setchellanthaceae	04	134213628-04			
Limnanthaceae	05	134213628-05			
Koerberliniaceae	06	134213628-06			
Bataceae	07	134213628-07			
Salvadoraceae	08	134213628-08			
Emblingiaceae	09	134213628-09			
Pentadiplandraceae	10	134213628-10			
Resedaceae	11	134213628-11			
Gyrostemonaceae	12	134213628-12			
Tovariaceae	13	134213628-13			
Capparaceae	14	134213628-14			
Cleomaceae	15	134213628-15			
Brassicaceae	16	134213628-16			
Santalales	11111111111111000000000000	134213632	Erythropalaceae	00	134213632-00
			Strombosiaceae	01	134213632-01
			Coulaceae	02	134213632-02
			Ximeniaceae	03	134213632-03
			Aptandraceae	04	134213632-04
			Olacaceae	05	134213632-05
			Octoknemaceae	06	134213632-06
			Schoepfiaceae	07	134213632-07
			Misodendraceae	08	134213632-08
			Loranthaceae	09	134213632-09
			Opiliaceae	10	134213632-10
			Santalaceae	11	134213632-11
Balanophoraceae	12	134213632-12			
Berberidopsidales	11111111111111000000000000	134215680	Aextoxicaceae	00	134215680-00
			Berberidopsidaceae	01	134215680-01



목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
Caryophyllales	1111111111111111110000000000	134216704	Droseraceae	00	134216704-00
			Nepenthaceae	01	134216704-01
			Drosophyllaceae	02	134216704-02
			Ancistrocladaceae	03	134216704-03
			Dioncophyllaceae	04	134216704-04
			Frankeniaceae	05	134216704-05
			Tamaricaceae	06	134216704-06
			Plumbaginaceae	07	134216704-07
			Polygonaceae	08	134216704-08
			Rhabdodendraceae	09	134216704-09
			Simmondsiaceae	10	134216704-10
			Asteropeiaceae	11	134216704-11
			Physenaceae	12	134216704-12
			Macarthuraceae	13	134216704-13
			Microteaceae	14	134216704-14
			Caryophyllaceae	15	134216704-15
			Achatocarpaceae	16	134216704-16
			Amaranthaceae	17	134216704-17
			Stenospermataceae	18	134216704-18
			Limeaceae	19	134216704-19
			Lophiocarpaceae	20	134216704-20
			Kewaceae	21	134216704-21
			Barbeuiaceae	22	134216704-22
			Aizoaceae	23	134216704-23
			Phytolaccaceae	24	134216704-24
			Gisekiaceae	25	134216704-25
			Sarcobataceae	26	134216704-26
			Nyctaginaceae	27	134216704-27
			Petiveriaceae	28	134216704-28
			Molluginaceae	29	134216704-29
			Montiaceae	30	134216704-30
			Halophytaceae	31	134216704-31
			Didiereaceae	32	134216704-32
			Basellaceae	33	134216704-33
			Talinaceae	34	134216704-34
			Anacampserotaceae	35	134216704-35
			Portulacaceae	36	134216704-36
Cactaceae	37	134216704-37			
Cornales	1111111111111111110000000000	134217216	Loasaceae	00	134217216-00
			Hydrangeaceae	01	134217216-01



목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
			Nyssaceae	02	134217216-02
			Cornaceae	03	134217216-03
			Curtisiaceae	04	134217216-04
			Grubbiaceae	05	134217216-05
			Hydrostachyaceae	06	134217216-06
Ericales	1111111111111111110000000	134217472	Balsaminaceae	00	134217472-00
			Marcgraviaceae	01	134217472-01
			Tetrameristaceae	02	134217472-02
			Polemoniaceae	03	134217472-03
			Fouquieriaceae	04	134217472-04
			Lecythidaceae	05	134217472-05
			Sladeniaceae	06	134217472-06
			Pentaphylacaceae	07	134217472-07
			Sapotaceae	08	134217472-08
			Ebenaceae	09	134217472-09
			Primulaceae	10	134217472-10
			Mitrastemonaceae	11	134217472-11
			Theaceae	12	134217472-12
			Symplocaceae	13	134217472-13
			Styracaceae	14	134217472-14
			Diapensiaceae	15	134217472-15
			Actinidiaceae	16	134217472-16
			Roridulaceae	17	134217472-17
			Sarraceniaceae	18	134217472-18
			Clethraceae	19	134217472-19
			Cyrillaceae	20	134217472-20
Ericaceae	21	134217472-21			
Garryales	1111111111111111110000000	134217600	Garryaceae	00	134217600-00
			Eucommiaceae	01	134217600-01
Gentianales	1111111111111111110100000	134217632	Rubiaceae	00	134217632-00
			Loganiaceae	01	134217632-01
			Gelsemiaceae	02	134217632-02
			Gentianaceae	03	134217632-03
Solanales	1111111111111111110110000	134217648	Apocynaceae	04	134217632-04
			Montiniaceae	00	134217648-00
			Sphenocleaceae	01	134217648-01
			Hydroleaceae	02	134217648-02
			Convolvulaceae	03	134217648-03
Boraginales	1111111111111111110111000	134217656	Solanaceae	04	134217648-04
			Codonaceae	00	134217656-00



목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
			Wellstediaceae	01	134217656-01
			Boraginaceae	02	134217656-02
			Hydrophyllaceae	03	134217656-03
			Namaceae	04	134217656-04
			Heliotropiaceae	05	134217656-05
			Cordiaceae	06	134217656-06
			Ehretiaceae	07	134217656-07
Lamiales	111111111111111111011100	134217660	Plocospermataceae	00	134217660-00
			Carlemanniaceae	01	134217660-01
			Oleaceae	02	134217660-02
			Tetrachondraceae	03	134217660-03
			Peltantheraceae	04	134217660-04
			Calceolariaceae	05	134217660-05
			Gesneriaceae	06	134217660-06
			Plantaginaceae	07	134217660-07
			Scrophulariaceae	08	134217660-08
			Stilbaceae	09	134217660-09
			Byblidaceae	10	134217660-10
			Linderniaceae	11	134217660-11
			Pedaliaceae	12	134217660-12
			Martyniaceae	13	134217660-13
			Acanthaceae	14	134217660-14
			Bignoniaceae	15	134217660-15
			Schlegeliaceae	16	134217660-16
			Lentibulariaceae	17	134217660-17
			Thomandersiaceae	18	134217660-18
			Verbenaceae	19	134217660-19
			Lamiaceae	20	134217660-20
			Mazaceae	21	134217660-21
			Phrymaceae	22	134217660-22
			Paulowniaceae	23	134217660-23
Orobanchaceae	24	134217660-24			
Aquifoliales	111111111111111111100000	134217664	Cardiopteridaceae	00	134217664-00
			Stemonuraceae	01	134217664-01
			Aquifoliaceae	02	134217664-02
			Helwingiaceae	03	134217664-03
			Phyllonomaceae	04	134217664-04
Asterales	111111111111111111100000	134217696	Rousseaceae	00	134217696-00
			Campanulaceae	01	134217696-01
			Pentaphragmataceae	02	134217696-02



목명	목 부호(2진수)	목 부호(10진수)	과명	순서	목-과 부호
			Alseuosmiaceae	03	134217696-03
			Phellinaceae	04	134217696-04
			Argophyllaceae	05	134217696-05
			Stylidiaceae	06	134217696-06
			Menyanthaceae	07	134217696-07
			Goodeniaceae	08	134217696-08
			Calyceraceae	09	134217696-09
			Asteraceae	10	134217696-10
Escalloniales	111111111111111111111111111111110000	134217712	Escalloniaceae	00	134217712-00
Bruniales	11111111111111111111111111111111000	134217720	Bruniaceae	00	134217720-00
			Columelliaceae	01	134217720-01
Apiales	1111111111111111111111111111111100	134217724	Pennantiaceae	00	134217724-00
			Torricelliaceae	01	134217724-01
			Griselinaceae	02	134217724-02
			Pittosporaceae	03	134217724-03
			Araliaceae	04	134217724-04
			Myodocarpaceae	05	134217724-05
			Apiaceae	06	134217724-06
Paracryphiales	111111111111111111111111111111110	134217726	Paracryphiaceae	00	134217726-00
Dipsacales	111111111111111111111111111111111	134217727	Viburnaceae	00	134217727-00
			Caprifoliaceae	01	134217727-01

## 2. 목-과 부호의 활용 방안

과명과 부호의 병기를 통해 해당 분류계급 간의 계통학적 상대적 거리를 직관적으로 알 수 있다. 예를 들어 장미과(Rosaceae)의 경우 과 부호는 '134213344-00', 두릅나무과(Araliaceae)와 산형과(Apiaceae)의 경우 과 부호는 각각 '134217724-04', '134217724-06'으로 표기된다. 이 경우 두릅나무과와 산형과는 동일 목에 속하므로 과 수준에서 가까운 근연 관계임을 유추할 수 있으며, 장미과와 이 두 개과는 목 부호 차이가 '4,380'으로 목 수준에서도 계통학적으로 상당히 차이 나는 그룹임을 알 수 있다. 또한, 국화과(Asteraceae, 134217696-10)와 꿀풀과(Lamiaceae, 134217660-20)의 목 부호 차이값은 '36'인데 비해 국화과와 작약과(Paeoniaceae, 134211584-01)의 목 부호 차이값은 '6,112'로, 국화과는 작약과보다는 꿀풀과에 가까운 관계임을 알 수 있다. 부호의 차이값은 상대적인 값으로 목 수준에서의 상대적 거리를 직관적으로 추론할 수 있으나, 차이값이 작으면 정확한 유연관계를 반영하기 어렵다는 단점이 있다. 이는 추후 연구 및 개정을 통해 가까운 계통을 지닌 분류군 간의 효율적인 부호 제시를 고민해야 할 것이다.

또한, 데이터베이스에 과 부호 필드[속성 예: CHAR(12)]를 추가하는 간단한 작업으로 여러 가지 활용이 가능하며, 계통에 따른 정렬 역시 가능하다. 데이터베이스에 단순히 과명만 있는 경우 알파벳순 정렬이 가능할 뿐이지만, 과 부호가 있을 때에는 양치식물(Pteridophyta)부터 진정국화군(Euasterids)



에 이르는 계통 순으로 정렬이 가능하다. 이러한 체계를 적용함으로써 웹 기반 각종 체크리스트 (checklist), 디지털플로라(digital flora) 및 전통의약정보체계의 정렬과 업데이트 시 작업 효율을 높여 준다. 또한 각 기관마다 구축하고 있는 데이터베이스에 이 부호를 적용하면 각 기관이 인식하고 있는 분류체계, 데이터베이스 간의 호환성 및 비교·대조에 효율적으로 접근할 수 있을 것이다.

더불어 목 단위에서 분계도의 시각화가 가능하다. 비록 목 이상의 분계군 명칭은 포함되지 않으나, 그림 1 과 같이 10 진수 부호를 2 진수로 변환하고 적절한 알고리즘을 거치면 화면에 그림 2와 같은 분계도를 자동으로 그려줄 수 있다. 트리 구조의 데이터 없이도, 특정한 분류군이 전체 관속식물 계통 중 어느 위치에 있는지를 시각적으로 확인할 수 있는 것이다. 또한 웹 상에서 비교 대상 분류군을 클릭 후 해당 분류군의 상대적 거리와 위치를 직관적으로 시각화할 수 있다.

```

$qq = array('Lycopodiaceae' => '000000000-00', 'Isoëtaceae' => '0335544');
$di = 27; // 목 부호 2진 자릿수
$q2 = array();
foreach($qq as $k => $v) {
    $v = explode('-', $v); // 목 부호와 과 부호 분할
    $q2[$v[0]][] = $k;
}
$q3 = array();
foreach($q2 as $k => $v) {
    $k = decbin($k); // 목 부호를 2진수로 변환
    $k = str_pad($k, $di, 0, STR_PAD_LEFT); // 자릿수 채우기
    $q3[$k] = implode(',', $v);
}
$q_ = array_reverse($q3); // 그리기 기호 변환을 위해 목 부호 배열 역전
$q4 = array();
$r = array();
$qi = 0;
foreach($q_ as $v => $k) {
    $o = str_split($v); // 목 부호 자릿수별 분리
    $o_ = array_reverse($o); // 그리기 기호 생성을 위해 순서 역전
    $o1 = array();
    $o1t = 0;
    for($i = 0; $i < count($o); $i++) {
        if($o_[$i] == 1 && $o1t != 1) {
            array_push($o1, 'L'); // 마지막 '1'은 'L' 기호로 변환
            $qt[$qi][] = 1;
            $o1t = 1;
        } elseif($o_[$i] == 1) {
            array_push($o1, ' '); // 나머지 '1'은 공백 기호로 변환
            $qt[$qi][] = 0;
        } elseif($o_[$i] == 0 && $o1t == 1) {
            array_push($o1, '|'); // 마지막 '1' 직전의 '0'은 '|' 기호로 변환
            $qt[$qi][] = 1;
        } elseif($o_[$i] == 0 && $qt[$qi-1][$i] == 1) {
            array_push($o1, '+'); // 아랫줄의 해당 자리가 'L'이나 '|'인 경우
            $qt[$qi][] = 0;
        } else {
            array_push($o1, '-'); // 나머지 '0'은 '-' 기호로 변환
            $qt[$qi][] = 0;
        }
    }
    $qi = $qi + 1;
    $o = array_reverse($o1); // 출력을 위해 그리기 기호 배열 역전
    $o = implode('', $o);
    $r[] = $o.' <span>'. $k.' </span>'; // 그리기 기호 뒤에 과명 나열
}
$r = array_reverse($r);
$r = implode('<br/>', $r);
print_r($r);

```

그림 1. 목-과 부호를 이용한 분계도 출력 프로그램 코드 예시





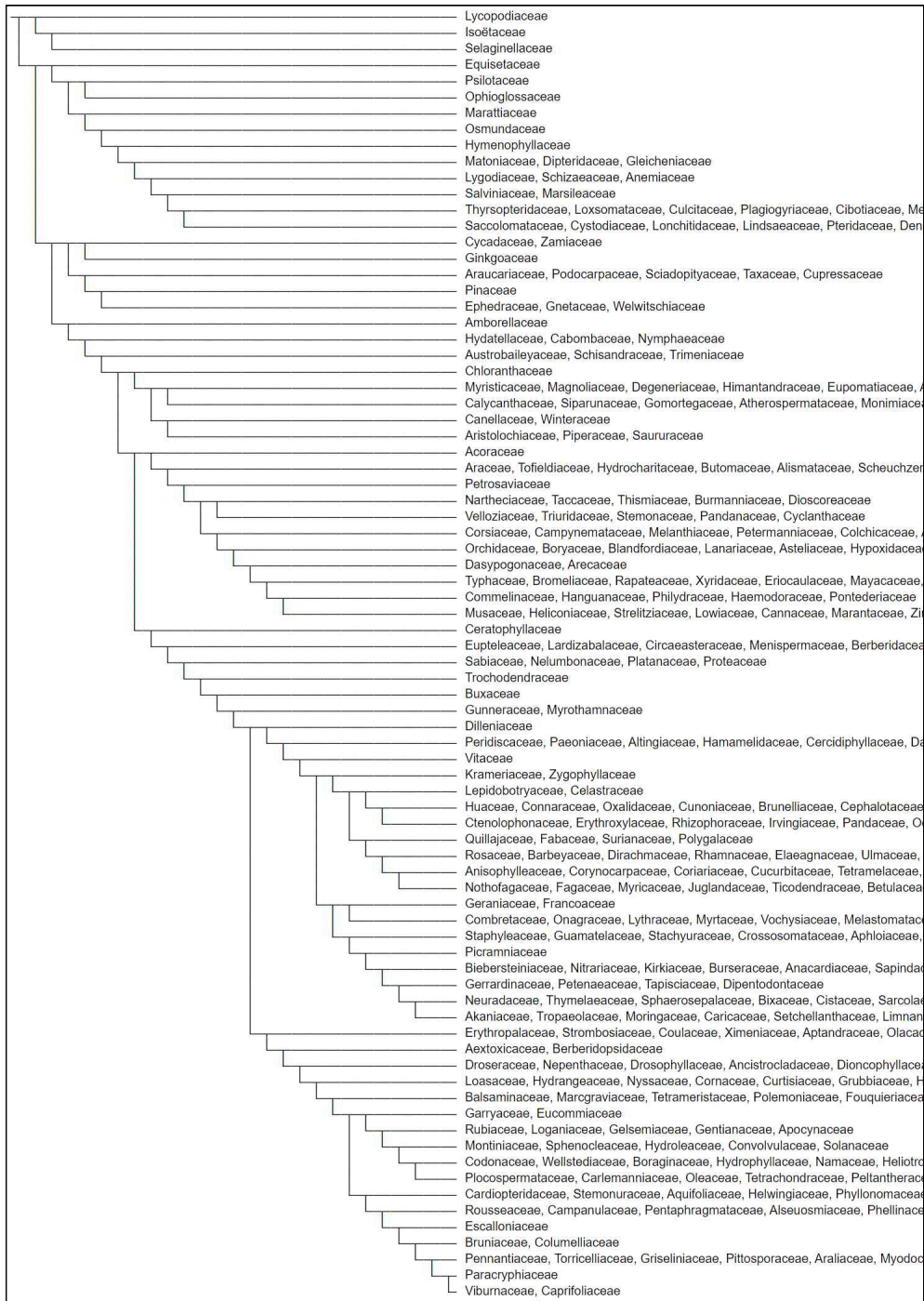


그림 2. 목-과 부호를 이용해 자동으로 생성된 분계도 예시('그림 1'의 코드 실행 화면)

### 3. 목-과 부호의 한계점 및 후속 연구 방향

목-과 부호는 정보 형태로서는 매우 간단하지만, 데이터베이스에 하나의 필드를 추가하는 작업은 곤란할 수 있다. 그리고, 목 내에서 과 계급의 위치 변동이 있을 경우에는 쉽게 수정할 수 있으나 새로운 목이 인식되어 추가되는 경우에는 전체 부호의 개정이 이루어져야 하므로 전면적인 수정이 불가피하다. 또한, 이 부호는 2020년 7월까지 AP Web에 제시된 연구결과만 반영된 부호이므로, 향후 계통학적 연구에 따라 목 간의 분지 위치 변동 또는 과 수준에서의 유연관계 재정립이 발생한 경우에는 분류체계의 개정에 따라 부호를 개정해야 한다는 단점이 있다. 아울러, 이 부호는 Angiosperm Phylogeny Group IV (APG IV)를 기반으로 한 AP Web의 체계를 이용한 것으로<sup>11,12)</sup> 이전의 앵글러(Engler) 분류체계<sup>14)</sup>, 크론크스트(Cronquist) 분류체계<sup>15)</sup>, 타크타잔(Takhtajan) 분류체계<sup>16)</sup> 등 다른 분류체계와는 호환되지 않는다.

본 부호는 비교적 상위 계급에 부여한 부호이므로, 과 이하 분류체계 적용을 위해서는 보다 효율적이고 직관적인 부호의 고민이 필요하다. 예컨대 《중국식물지》의 Hu Card Index<sup>17)</sup>는 중국의 학자 Hu Shiu-ying과 연구진이 중국에서 출판된 식물 학명을 수집하여 정리한 내용으로 종 하 분류군을 포함한 중국식물 전체 학명이 망라되어 있다(1753년~1955년 출판물 기준). 웹 사이트에는 과명과 속명에 대한 코드(숫자)가 부여되어 있으나 과를 기반으로 속과 종까지 코드가 부여된 것으로, 동일 과 내에서도 속과 종이 다수인 경우 코드의 폭이 큰 것으로 확인되었다. 예를 들어 벼과(Poaceae)의 경우 Hu Card Index는 '004345'에서 '010212'까지 부여되어 있으며, 종 이하 분류계급의 경우에는 따로 코드가 부여되지 않아 상대적 거리 유추 및 분류군의 효율적인 정렬이 어려운 것으로 나타났다. 다만, Hu Card Index 코드(number)와 Tropicos ID 및 과명(family name)과의 연계를 위한 비교·제시를 통해 타 정보체계(Tropicos)와의 호환성 문제가 중요함을 확인할 수 있었다. 속이나 종 수준에서 동속이종의 학명 제시, 주요 기원 혼란종, 유사종 등의 연관 분류군 제시 등 확장성이 필요한 경우 Hu Card Index와 같이 기존에 활용되고 있는 다양한 부호, 코드, 인덱스를 비교·검토할 필요가 있다.

데이터베이스 구축을 위한 목적성을 분명히 하고, 구축할 전체 대상의 범주를 정확히 파악한 후 가장 효율적인 부호를 부여한다면 단순한 데이터베이스의 구축뿐만 아니라 대상 품목·분류군의 표본·분포 관리, 연관 문헌 및 도출 데이터 관리를 포함한 다양한 데이터를 연계하여 종합적으로 제시할 수 있을 것이다.

## 결론

생물정보학 내 생물정보 관리의 관점에서 관속식물 분류체계 데이터베이스 구축을 위한 '목-과 부호'(order-family code)를 제안하였다.

1. 각 과 사이의 유연관계를 포함하면서 데이터베이스 자원도 적게 소비하는 10진수 12자리 목-과 부호를 부여하였다.
2. 과명과 부호의 병기를 통해 해당 분류계급 간의 계통학적 상대적 거리를 직관적으로 알 수 있으며, 데이터베이스에 과 부호 필드 추가, 분계도의 시각화 등 다양한 방법으로 적용이 가능하다.
3. 추후 과 이하 속, 종 수준에서 데이터베이스의 확장성이 필요한 경우 추가적인 후속 연구의 방향성을 제안하였다.



## 감사의 글

본 연구는 ‘2020년도 식품의약품안전처의 연구개발비(과제번호: 18172 생약안 201)’로 수행되었으며 이에 감사드립니다.

## 참고문헌

1. Robbins RJ. Bioinformatics: essential infrastructure for global biology. *Journal of Computational Biology*. 1996;3(3):465-78.
2. Rhee SY, Dickerson J and Xu D. Bioinformatics and its applications in plant biology. *Annual Review of Plant Biology*. 2006;57:335-60.
3. Zhong Y, Luo Y, Pramanik S and Beaman JH. HICLAS: a taxonomic database system for displaying and comparing biological classification and phylogenetic trees. *Bioinformatics (Oxford, England)*. 1999;15(2):149-56.
4. Integrated Taxonomic Information System (ITIS). Available from: <http://www.itis.gov> (accessed 2020-05-27).
5. Royal Botanic Gardens, Kew Science. International Plant Names Index (IPNI). Available from: <https://www.ipni.org> (accessed 2020-05-27).
6. The Plant List. Available from: <http://www.theplantlist.org> (accessed 2020-05-27).
7. Index Fungorum. Available from: <http://www.indexfungorum.org> (accessed 2020-05-27).
8. The Marine Biological Laboratory. Universal Biological Indexer and Organizer (uBio). Available from: <http://www.ubio.org> (accessed 2020-05-27).
9. National Center for Biotechnology Information, U.S. National Library of Medicine. GenBank. Available from: <http://www.ncbi.nlm.nih.gov/Genbank/GenbankOverview.html> (accessed 2020-05-27).
10. The Barcode of Life Data System. Available from: <http://www.boldsystems.org/index.php> (accessed 2020-05-27).
11. The Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*. 2016;181(1):1-20.
12. Stevens, P. F. (2001 onwards). Angiosperm Phylogeny Website. Version 14. Available from: <http://www.mobot.org/MOBOT/research/Apweb/> (accessed 2020-07-03).
13. The Pteridophyte Phylogeny Group. A community-derived classification for extant lycophytes and ferns. *Journal of Systematics and Evolution*. 2016;54(6):563-603.
14. Engler A. *Führer durch den Königlich botanischen garten der universität zu Breslau*. Breslau:Botanischer Garten. 1886.
15. Cronquist A and Takhtajan A. *An Integrated System of Classification of Flowering Plants*. New York:Columbia University Press. 1981.

16. Takhtajan A. Diversity and Classification of Flowering Plants. New York:Columbia University Press. 1997.
17. Missouri Botanical Garden. Tropicos.org., Release 3.0.2. Available from: <http://www.tropicos.org> (accessed 2020-05-27).

© The Author(s) 2020, khmi.or.kr

